Quality of life assessment in gastro-oesophageal reflux disease

E J Irvine

Health related quality of life (HRQoL) is determined by both disease and non-disease related factors. Several studies have reported significant HRQoL impairment in GORD patients compared with the general population. Disease severity correlates strongly with HRQoL. Non-disease features, such as the presence of anxiety and comorbid conditions, also negatively impact on HRQoL. Combining a generic and disease specific instrument may avoid missing unexpected outcomes and ensure recognition of all clinically important changes. Full validation of assessment tools is critical. Long term, as well as short term, evaluation is important and is critical when undertaking comparative pharmacoeconomic evaluations.

SUMMARY
Health related quality of life (HRQoL) is determined by both disease and non-disease related factors. In chronic illnesses, such as gastro-oesophageal reflux disease (GORD), daily function, HRQoL status, and health resource utilisation are critical outcomes. Patients suffering from GORD report many symptoms, such as heartburn or regurgitation, and health care seeking is driven by both symptom severity and the impact on HRQoL. Some individuals intuitively alter their lifestyles while others do not. Most who have bothersome symptoms desire effective medical or surgical treatment. Several descriptive studies have reported significant HRQoL impairment in GORD patients compared with the general population, similar to other chronic conditions, such as myocardial infarction. Disease severity correlates strongly with HRQoL and also contributes to work absenteeism and reduced productivity. Non-disease features, such as the presence of anxiety and comorbid conditions, also negatively impact on HRQoL. Several generic and disease specific HRQoL instruments have been applied in patients with GORD. In clinical trials, the psychological general well being index (PGWBI), SF-36, quality of life in reflux and dyspepsia (QOLRAD) scale, and gastrointestinal quality of life index (GIQLI) have assessed HRQoL as a secondary outcome in patients with GORD. The degree of validation and psychometric assessment of some questionnaires varies. Few have been sufficiently assessed to fully recommend their use. Selecting the most appropriate HRQoL instrument should be based on the research question and study design. However, psychometric robustness is critical to accurately interpret results. Combining a generic and disease specific instrument may avoid missing unexpected outcomes and ensure recognition of all clinically important changes. Effective GORD treatment rapidly improves HRQoL but long term outcomes are most important for this chronic health problem, and must be incorporated into economic evaluations.

QUALITY OF LIFE
Evaluating the impact of GORD has relied heavily in the past on scales assessing symptom severity, such as heartburn, regurgitation, or pain, together with the endoscopic appearance of the oesophageal mucosa. Laboratory markers such as anemia are too insensitive for diagnosis, except in the most severe GORD cases, and pH studies, although useful in selected subjects, are generally not used to assess therapeutic response. Moreover, these assessment methods, even when used collectively, still fail to reflect the functional status of some patients. HRQoL has surfaced therefore as a clinically relevant measure of disease impact and treatment response. HRQoL can be defined as the functional effect of an illness and its therapy on an individual, as perceived by the individual himself or herself. The domains that determine HRQoL include physical and occupational function, emotional state, social interactions, and somatic sensation. These determinants can also be further classified as simply disease related, including symptom severity, treatment efficacy, and adverse effects of treatment, or disease independent factors, such as sex or age, education and knowledge, personality and coping skills, culture, and beliefs. A list of the relevant domains and items that measure HRQoL in patients with GORD are listed in table 1.

Several decades ago, Engel proposed a biopsychosocial model of chronic disease that has been appropriately applied in gastrointestinal disorders such as GORD or irritable bowel syndrome. Briefly, the model acknowledges that an individual may have a genetic or biological predisposition to a condition, such as GORD, and that

Correspondence to: Professor E J Irvine, Department of Medicine, University of Toronto, Head, Division of Gastroenterology, St Michael’s Hospital, Room 16-052 Cardinal Carter Wing, 30 Bond St, Toronto, ON, M5B 1W8, Canada; irvinej@smh.toronto.on.ca

Abbreviations: HRQoL, health related quality of life; gastro-oesophageal reflux disease; PGWBI, psychological general well being index; QOLRAD, quality of life in reflux and dyspepsia; GIQLI, gastrointestinal quality of life index; SF-36, medical component summary score; GQARY, quality adjusted life year; HUI, health utilities index mark III; GSRS, gastrointestinal symptom rating scale; PCS score, physical component summary score; MCS score, mental component summary score
and liver disease (reviewed by Borgaonkar and Irvine) and available describing HRQoL assessment in gastrointestinal measures, such as the SF-36 or PGWBI, allow comparisons of determinants of good or poor HRQoL. While generic summary of function but fail to identify the important risks or benefits. A considerable literature is scored by examining individual items such as heartburn, regurgitation, or pain, clusters of similar items in subscores (such as vitality or emotional function), or by summing all the items into a global or summary score. These different scoring methods allow identification of common problems of individuals, groups, or populations, allow comparisons at a single time point, or help measure change over time (natural history) or after therapy (clinical trials). HRQoL assessment thus provides a useful yardstick for patients attempting to improve their HRQoL, and for clinicians, researchers, or policy makers by helping identify the needs of individual patients, assessing the impact of therapy (in individuals or in clinical trials), and determining health policy. In general, improving the patient’s HRQoL is a common objective that can be effected using a variety of approaches. HRQoL assessment can also be applied in pharmacoeconomic analyses when comparing costs to achieve a particular outcome (such as remission, no symptoms), or costs per quality adjusted life year (QALY) gained. Utilities are patient generated preference “weights”, ranging from 0.0 (death) to 1.0 (full health), that are often derived in a reference population using instruments such as the time trade off or the health utilities index mark III (HUI). One QALY represents one full year in perfect health. Different utility instruments such as different HRQoL tools may elicit different values in the same reference population but may well produce similar results when measuring changes over time or after treatment.

HRQOL MEASUREMENT
Evaluation of HRQoL is conventionally performed using questionnaires or surveys that can be scored quantitatively and address problems that occur as a consequence of the chronic illness, as well as features of psychosocial well being. The three types of instrument that measure HRQoL (reviewed by Guyatt and colleagues) include global assessments (single descriptors), generic measures (multi-item profiles generated in a general population that can be applied to different diseases), and disease specific instruments (developed in and for patients affected by a common condition). Global assessments are simple to administer and provide a useful summary of function but fail to identify the important determinants of good or poor HRQoL. While generic measures, such as the SF-36 or PGWBI, allow comparisons of groups with different conditions or help identify unexpected HRQoL issues, it is the disease specific questionnaire that is most likely to detect small but clinically important risks or benefits.

A considerable literature is available describing HRQoL assessment in gastrointestinal and liver disease (reviewed by Borgaonkar and Irvine) and attributes and review criteria to select the most appropriate instrument (reviewed by the Scientific Advisory Committee of the Medical Outcomes Trust). Health status scales, such as HRQoL indices, can be scored by examining individual items such as heartburn, regurgitation, or pain, clusters of similar items in subscores (such as vitality or emotional function), or by summing all the items into a global or summary score. These different scoring methods allow identification of common problems of individuals, groups, or populations, allow comparisons at a single time point, or help measure change over time (natural history) or after therapy (clinical trials). HRQoL assessment thus provides a useful yardstick for patients attempting to improve their HRQoL, and for clinicians, researchers, or policy makers by helping identify the needs of individual patients, assessing the impact of therapy (in individuals or in clinical trials), and determining health policy. In general, improving the patient’s HRQoL is a common objective that can be effected using a variety of approaches. HRQoL assessment can also be applied in pharmacoeconomic analyses when comparing costs to achieve a particular outcome (such as remission, no symptoms), or costs per quality adjusted life year (QALY) gained. Utilities are patient generated preference “weights”, ranging from 0.0 (death) to 1.0 (full health), that are often derived in a reference population using instruments such as the time trade off or the health utilities index mark III (HUI). One QALY represents one full year in perfect health. Different utility instruments such as different HRQoL tools may elicit different values in the same reference population but may well produce similar results when measuring changes over time or after treatment.

It is the responsibility of the research team to ensure that the HRQoL or utility instrument used for any given study clearly addresses the problems of the research question. Most importantly, it should have undergone a full psychometric assessment (with respect to validity, reliability, and responsiveness). "Validity" is a comparison of the new index score and a reference score (convergent validity), or a construct of what the new index is measuring, such as a
prediction that patients with more severe disease will have poorer HRQoL scores (construct validity). “Reliability” is an assessment of the measurement error of scores (test–retest reliability) or the correlation among items or subscores (internal reliability). “Responsiveness” gives a signal to noise ratio and allows interpretation of what degree of change is clinically important. In selecting a questionnaire, it is important that it performs robustly in the population being sampled and for the study design planned. The results must be analysed according to the research methods and interpreted objectively, addressing any biases that may confound the results.

**PROBLEMS OF GORD PATIENTS AND DETERMINANTS OF HRQOL**

Heartburn is the most common symptom, identified by 89% of GORD patients, and the most prevalent symptom required for eligibility in clinical trials. Regurgitation and heartburn are also assessed in the gastrointestinal symptom rating scale (GSRS). Other symptoms affecting at least 25% of subjects with GORD include belching, flatulence, acid taste, nausea, stomach gurgling, and bad breath. Symptoms are an important part of outcome and may be included as part of a disease specific instrument or separately as a disease severity index.

Few studies have assessed the relative importance of symptoms and other HRQoL problems in patients with GORD. Talley and colleagues, in assessing 984 patients entering two clinical trials with endoscope negative reflux disease, observed that the most commonly reported areas of HRQoL dysfunction were eating and drinking (45–81%), sleep problems (39–49%), lack of vitality (41–58%), and poor emotional well being (45–55%). Most daily activities were impacted in approximately 20% of subjects, and as many as 44% avoided bending over because of heartburn. Somewhat different symptoms occur in patients who have undergone laparoscopic or open surgery who report flatulence (up to 40%), bloating, and postprandial fullness (10–20%), and dysphagia (10%) after surgery but with no apparent differences between the two types of operation. Patient dysphagia is also apparent in the workplace. In a prospective descriptive study of 136 Swedish workers attending their general practitioner for GORD, 30% reported decreased regular daily activities, 23% had reduced weekly sleeping, vitality, or emotional well being. Within two weeks of treatment with esomeprazole, 20 or 40 mg daily, these subjects, 40–80% reported problems in eating/drinking, sleeping, vitality, or emotional well being. Within two weeks of treatment with omeprazole or esomeprazole were less likely to respond to therapy than groups reporting little or no anxiety. These results support the hypothesis that symptom severity and non-disease related factors, such as sex, age, comorbidity, and personality traits also contribute to HRQoL outcomes.

There are now at least three generic and six disease specific instruments that have been evaluated in patients with GORD (table 2). Some of these have been well evaluated while others have not. The two most extensively examined are the GRS and the QOLRAD scale. These have been fully evaluated and demonstrated to be valid, reliable, and sensitive to change. None the less, the GRS is primarily a symptom severity score rather than a true HRQoL instrument and thus should be considered separately and not as a quality of life instrument. Talley and colleagues fully assessed psychometrically the QOLRAD using results obtained during two randomised controlled trials that examined the GRS and QOLRAD in 984 patients. Of these subjects, 40–80% reported problems in eating/drinking, sleeping, vitality, or emotional well being. Within two weeks of treatment with esomeprazole, 20 or 40 mg daily, these problems had decreased to 7–19%, and there was an excellent correlation between symptom reduction and improved function. The minimum clinically important difference was identified to be approximately 0.5 per item, with a large difference being approximately 1.5 points per item.

The effects of treatment on HRQoL have been reported in several observational studies but only a few randomised trials. One systematic review was published in the Cochrane database in 2000 reporting medical therapy in 23 trials, of which only four had included HRQoL as a measured outcome. Three used the GRS, a symptom scale, two also used the PGWBI, and one used the SF-36 together with a heartburn specific instrument. Significantly improved PGWBI scores and GRS reflex scores (p<0.05) were noted after treatment of endoscope negative reflex disease with omeprazole 20 or 10 mg daily compared with placebo. No overall difference was noted between omeprazole 20 mg once
daily and cisapride 10 mg four times daily, except in the GSRS reflux score, and significant improvement was noted also with ranitidine 150 mg twice daily compared with placebo in three (physical functioning, bodily pain, and vitality) of eight subscales of the SF-36 and all scales of the heartburn specific questionnaire. Currently, a second overview is in progress and has reviewed a total of 174 short and long term GORD trials, of which only nine have assessed HRQoL (Moayyedi P, personal communication).

In a retrospective comparison of symptoms experienced over the previous 12 months by patients with severe reflux who had undergone medical or surgical management, heartburn, regurgitation, waterbrash, and HRQoL were significantly better (p<0.05) after laparoscopic fundoplication compared with medical management. As might be expected, medication and proton pump inhibitor use was greater in the medically treated group and no disease specific HRQoL was assessed. However, the risk of death was greater in the surgically treated patients but did show greater bodily pain in the medical treatment group (51.7 v 64.0; p<0.05). However, the risk of death was greater in the surgically treated group and no disease specific HRQoL was assessed.

Most studies in which effective therapy has been given, whether medical or surgical, show improvement in HRQoL after treatment. It seems important to remind researchers that when assessing HRQoL, all important outcomes must be determined. Finally, in recent times, studies are undertaken in many countries simultaneously. It is critical to ensure that proper cross cultural validation is performed when HRQoL questionnaires must be translated or adapted for other languages.

Most researchers are now combining generic and disease specific instruments to fully assess HRQoL. This has the advantage of not missing an unexpected finding, and also of finding the most important benefit. However, alternative instruments will be needed for broader or more specific applications, such as to assess mental health, work function, disease knowledge, coping skills, relationships, or sexual dysfunction, which are not fully assessed by currently available questionnaires. Questionnaires, such as the World Health Organisation quality of life (WHOQOL-100), a reasonably well validated but rather lengthy instrument (or its shorter version, the WHOQOL-BREF) have not been examined in the context of GORD but should be considered for evaluation. Utilities, such as the HUI3 and others, should also be explored when pharmacoeconomic analyses are planned.

In summary, GORD patients have impaired HRQoL compared with general populations, and perceive themselves to be as affected by their condition as groups with other serious chronic conditions. The level of impairment and types of problems experienced relate to symptom severity, the type and effectiveness of the treatment, and non-disease related factors such as the presence of other medical conditions, sex, or anxiety. HRQoL measurement complements symptom severity evaluation and should be part of outcome measurement for all therapeutic trials, using both disease specific and generic instruments. Full validation of assessment tools is critical. Long term, as well as short term, evaluation is important and is critical when undertaking comparative pharmacoeconomic evaluations.

### Table 2  Health related quality of life (HRQoL) instruments used in gastro-oesophageal reflux disease (GORD)

<table>
<thead>
<tr>
<th>Instrument</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Generic instruments used in the assessment of GORD</strong></td>
<td></td>
</tr>
<tr>
<td>SF-36</td>
<td>36 items; 8 subscores: physical function, social function, role limitation physical, role limitation emotional, bodily pain, mental health, vitality, health perception Score 0 (worst)–100 (best)</td>
</tr>
<tr>
<td>PGWB1</td>
<td>22 items; 6 subscales: anxiety, depression, well being, self control, general health, vitality, and energy Score 22–132, higher score better</td>
</tr>
<tr>
<td>Euro-QOL90</td>
<td>5 or 7 items utility; score 0 to 100 (best)</td>
</tr>
<tr>
<td><strong>Disease specific instruments developed for patients with GORD</strong></td>
<td></td>
</tr>
<tr>
<td>GSRS64</td>
<td>15 items, 5 clusters: reflux, abdominal pain, indigestion, diarrhoea, constipation More accurately a symptom rating scale 7 worst to 0 best</td>
</tr>
<tr>
<td>GERQ60</td>
<td>76–80 items; reflux symptoms plus SF-36. Heartburn, regurgitation, effect of heartburn, pain, dysphagia, UGI, respiratory past history, medications, port treatments, miscellaneous Heartburn scoring unclear</td>
</tr>
<tr>
<td>GORD-HRQoL65</td>
<td>10 items; 45 worst to 0 best</td>
</tr>
<tr>
<td>HBQOL32</td>
<td>15 items; 0 (worst)–100 (best) Role physical, pain, sleep, diet, social, mental health</td>
</tr>
<tr>
<td>QOLRAD33</td>
<td>25 items; 1 (worst) to 7 (best) Emotional, sleep, eating problems, physical/social, vitality</td>
</tr>
<tr>
<td>GIQLI63</td>
<td>Up to 36 items; 0 (worst) to 144 (best). Modular pain, gas, eating, mood, sleep, stamina, work, recreation, relationship, sex, oesophageal, colonic</td>
</tr>
</tbody>
</table>
REFERENCES

18 Talley NJ, Junghard O, Wicklund IK. Why do patients with gastroesophageal reflux disease have a poor health-related quality of life? Gastroenterology 2001;120:A423.